# Lassen sich mit SPSSx-Matrix anwenderspezifische Analyseprobleme lösen? Ein Anwendungstest am Beispiel der multinomialen logistischen Regression

Kühnel, Steffen M.

# Lassen sich mit SPSS$^x$-Matrix anwenderspezifische Analyseprobleme lösen? Ein Anwendungstest am Beispiel der multinomialen logistischen Regression[1]

von Steffen M. Kühnel

Seit Version 4.0 ist in SPSS$^X$ mit der Prozedur MATRIX die Möglichkeit gegeben, Analysemodelle über Matrizen-Operationen selber zu programmieren. Im Prinzip sollte es somit auch möglich sein, innerhalb des SPSS$^x$-Systems statistische Modelle anzuwenden, für die noch keine Standard-Prozedur existiert. Für eine übej didaktische Zwecke hinausgehende Nutzung ist es wichtig, daß solche "selbstgestrickten" Analyse-Programme auch größere Datenmengen in akzeptabler Zeit verarbeiten können.

Ob die SPSS$^x$-Prozedur MATRIX diesem Anspruch genügen kann, soll in diesem Beitrag an Hand eines Beispiels untersucht werden. Mein Anwendungsbeispiel ist die Implementation des Newton-Raphson-Algorithmus zur Maximum-Likelihood-Schätzung der Regressionskoeffizienten der multinomialen logistischen Regression. Das statistische Modell der logistischen Regression ist in den letzten Ausgaben der ZA-Information in mehreren Beiträgen behandelt worden (vgl. *Ludwig-Mayerhof er,* 1990; *Kühnel u.a.,* 1989; *Urban,* 1990). SPSS$^x$-Benutzer können jedoch bislang nur den Spezialfall der binären logistischen Regression anwenden. Eine Implementation des allgemeineren multinomialen Modells über die Prozedur MATRIX könnte den Anwenderkreis dieser Modelle möglicherweise erweitern.

Der vorüegende Beitrag ist so aufgebaut, daß die drei Hauptteile auch unabhängig voneinander gelesen werden können. Im ersten Teil gehe ich noch einmal knapp und recht formal auf das statistische Modell der logistischen Regression ein und beschreibe den verwendeten Algorithmus zur Schätzung der Modellparameter. Im zweiten Abschnitt wird über die Erfahrungen bei der Implementation des Modells berichtet.[2] Dabei gehe ich insbesondere auf Probleme ein, die möglicherweise auch bei anderen Anwendungsversuchen auftreten können. Im dritten Teil wird die praktische Anwendbarkeit des Matrixprogramms an einer sozialwissenschaftlichen Fragestellung geprüft. Inhaltlich geht es hier um die Fra-

---

1 Ich danke *Michael Terwey,* daß er das hier beschriebene Matrixprogramm aus der Sicht eines externen Anwenders testete und wertvolle Hinweise zur Verbesserung gab.

2 Das resultierende Matrixprogramm ist in ein SPSS$^x$-Makro eingebettet, das Interessenten auf Anforderung zur Verfügung gestellt wird. In Anhang 1 wird der Aufruf dieses Makros beschrieben.

ge, ob entsprechend der Handlungstheorie von ***Fishbein und Ajzen*** die Handlungsabsicht der beste und letzlich einzige Prädiktor von Handlungen ist.

## 1. Der Newton-Raphson-Algorithmus zur ML-Schätzung der Parameter eines multinomialen logistischen Regressionsmodells

Die logistische Regression ist ein Analysemodell zur Untersuchung des Einflusses einer oder mehrerer (metrischer) erklärender Variablen auf eine abhängige nominalskalierte Variable. Der Zusammenhang zwischen der abhängigen Variable und den erklärenden Variablen wird dabei so modelliert, daß das Verhältnis der Auftretenswahrscheinlichkeiten von jeweils zwei verschiedenen Kategorien der abhängigen Variable exponential mit einer Linearkombination der erklärenden Variablen variiert. $X_1, X_2, ..., X_K$ seien die insgesamt K erklärenden Variablen und Y die abhängige Variable mit den Kategorien i = 1, 2, ..., I. Für jeden Fall t einer Stichprobe von n Fällen wird im logistischen Regressionsmodell also unterstellt, daß für zwei Kategorien i und j von Y gilt:

$$(1) \qquad \frac{P(y_t=i)}{P(y_t=j)} = \exp\left(\sum_{k=1}^{K} \beta_{ijk} * x_{kt}\right) = e^{x_t'\beta_{ij}}$$

Von den nach Gleichung (1) theoretisch möglichen K*I*(I-1) Regressionskoeffizienten sind tatsächlich nur K*(I-1) nicht redundant. Vertauscht man nämlich in Gleichung (1) Nenner und Zähler, wird deutlich, daß gilt:

$$(2) \qquad \beta_{ijk} = -\beta_{jik}$$

Darüber hinaus gilt auch für beliebige Kategorien i,j und r der abhängigen Variable:

$$(3) \qquad \beta_{ijk} = \beta_{irk} - \beta_{jrk}$$

Letzteres folgt aus:

$$\text{(4)} \qquad \frac{P(Y_t=i)}{P(Y_t=j)} = \frac{\left(\frac{P(Y_t=i)}{P(Y_t=r)}\right)}{\left(\frac{P(Y_t=j)}{P(Y_t=r)}\right)}$$

Bei der Modellschätzung eines logistischen Regressionsmodells werden daher nur die K*(I-1) Regressionskoeffizienten $ß_{irk}$ einer Referenzkategorie r der abhängigen Variable geschätzt.[3] Die Schätzung dieser Modellparameter kann nach der Maximum-Likelihood-Methode erfolgen. Dabei werden die Koeffizienten $ß_{irk}$ so bestimmt, daß die Wahrscheinlichkeit maximal ist, gerade die in der Stichprobe aufgefundenen Werte der abhängigen Variable zu erhalten. Bei korrekter Modellspezifikation und einer einfachen Zufallsauswahl[4] voneinander unabhängiger Realisationen läßt sich die Likelihood-Funktion folgendermaßen darstellen:

$$\text{(5)} \qquad L = \prod_{t=1}^{n} \prod_{i=1}^{I} P(Y_t=i)^{d_{it}}$$

In (5) hat die Indikatorvariable $d_{it}$ den Wert Eins, wenn $p(y_t=i)$, ansonsten den Wert Null. Aus Gleichung 1 und der Summierung aller Wahrscheinlichkeiten zur Summe 1 folgt für die einzelnen Wahrscheinlichkeiten:

$$\text{(6)} \qquad \begin{aligned} P(Y_t=i) &= e^{x_t'\beta_{ir}} * P(Y_t=r); \qquad i \neq r \\ P(Y_t=r) &= \frac{1}{1+\sum_{i=1 \neq r}^{I} e^{x_t'\beta_{ir}}} \end{aligned}$$

---

3 Der Index r für die Referenzkategorie wird dabei oft ausgelassen. Bei der Interpretation der Koeffizienten muß natürlich die jeweils gewählte Referenzkategorie berücksichtigt werden.

4 Die folgende Funktion läßt sich auch bei geschichteten Stichproben anwenden. Bei endogen geschichteten Stichproben, bei denen die Auswahl über Teilstichproben für die einzelnen Kategorien der abhängigen Variable erfolgt, werden die Regressionskonstanten allerdings nicht konsistent geschätzt (vgl. *Maier u. Weiss,* 1990:207ff.).

Das Einsetzen von (6) in (5) zeigt, daß die Likelihood-Funktion tatsächlich eine mathematische Funktion der Regressionskoeffizienten $\beta_{irk}$ ist. Statt nun direkt das Maximum dieser Funktion zu bestimmen, wird das Negative des natürlichen Logarithmus der Funktion minimiert:

$$(7) \qquad -\ln L = -\sum_{t=1}^{n}\sum_{i=1}^{I} d_{it}\cdot\ln\left(P(y_t=i)\right) = -\sum_{t=1}^{n} d_t'\cdot\ln(p_t)$$

Bei unterstellter korrekter Modellspezifikation sind die so geschätzten Regressionskoeffizienten asymptotisch um die gesuchten Populationswerte multinormalverteilt. Über die ebenfalls schätzbaren Standardfehler der Schätzer können daher asymptotisch gültige Signifikanztests durchgeführt werden, die dem T-Test in der linearen Regression entsprechen. Darüber hinaus lassen sich auch asymptotisch gültige Chiquadrat-Tests von beliebigen Linearkombinationen der Regressionskoeffizienten berechnen, so daß sowohl Kontraste als auch der Einfluß jeder einzelnen erklärenden Variablen über die Gesamtheit der ihr zugeordneten (1-1) Koeffizienten getestet werden können. Schließlich sind auch Chiquadrat-Differenzentests (Likelihood-Ratio-Tests) hierarchisch geschachtelter Regressionsmodelle möglich. Unterscheiden sich nämlich zwei logistische Regressionsmodelle nur dadurch, daß in einem der beiden Modelle einige Regressionskoeffizienten a priori auf Null fixiert sind, so ist 2 mal die Differenz der Maxima der beiden negativen logarithmierten Likelihood-Funktionen (7) asymptotisch chiquadratverteilt. Den Chiquadrat-Differenzentests entsprechen in der linearen Regression die F-Tests von $R^2$ bzw. Änderungen von $R^2$. Ein dem Determinationskoeffizienten $R^2$ in gewisser Hinsicht analoges PRE-Maß zur Stärke des Zusammenhangs läßt sich durch den Vergleich der Minima von Gleichung (7) bei zwei hierarchisch geschachtelten Modellen konstruieren. Wenn $-\ln L_0$ das Minimum der Gleichung (7) eines Modells ist, das nur Regressionskonstanten enthält, und $-\ln L_1$ das Minimum eines Modells, das zusätzlich erklärende Variablen aufweist, berechnet sich dieses Pseudo-$R^2$ nach:

$$(8) \qquad P^2 = 1 - \frac{-\ln L_1}{-\ln L_0}$$

Die rechnerische Bestimmung der Maximum-Likelihood-Lösung kann nach dem Newton-Raphson-Algorithmus erfolgen.[5] Dazu benötigt man die ersten und zweiten Ableitungen der Funktion (7) nach den unbekannten Parameterwerten. Wenn I die Referenzkategorie ist und $\mathbf{B} = (\beta_{1I}, \beta_{2I},..., \beta_{I-1,I})$ die Matrix der Regressionskoeffizient $\mathbf{d}_t^*$ die ersten (I-1) Elemente des Vektors $\mathbf{d}_t$ aus Gleichung (7) enthält und $\mathbf{p}_t^*$ die entsprechenden Elemente von $\mathbf{p}_t$, dann ergeben sich die ersten Ableitungen als:

(9)
$$\frac{\partial \ln L}{\partial \boldsymbol{B}} = \sum_{t=1}^{n} \boldsymbol{x}_t(\boldsymbol{d}_t^* - \boldsymbol{p}_t^*)'$$

Die Matrix der zweiten Ableitungen berechnet sich nach:

(10)
$$\frac{\partial^2 \ln L}{\partial \mathrm{vec}(\boldsymbol{B})\partial \mathrm{vec}(\boldsymbol{B})'} = -\sum_{t=1}^{n} (\mathrm{diag}(\boldsymbol{p}_t^*) - \boldsymbol{p}_t^* \boldsymbol{p}_t^{*\prime}) \otimes (\boldsymbol{x}_t \boldsymbol{x}_t')$$

In Gleichung (10) steht "vec(.)" für die Umgruppierung der Elemente einer Matrix in einen Spaltenvektor und "diag(.)" für eine Diagonalmatrix, deren Diagonale die Elemente des in der Klammer aufgeführten Vektors enthalten. Das Symbol $\otimes$ steht für die Bildung des Kronekerprodukts zweier Matrizen.

Der Algorithmus führt zu einer schrittweisen Minimierung von Gleichung (7). Dabei werden in jedem Iterationsschritt s+1 die Koeffizienten des vorherigen Schrittes s korrigiert. Die Korrekturgröße ist dabei das Negative des Produkts der Inversen der Matrix der zweiten Ableitungen mit dem Vektor der ersten Ableitungen. Es gilt also:

(11)
$$\mathrm{vec}(\boldsymbol{B}^{s+1}) = \mathrm{vec}(\boldsymbol{B}^s) - \left(\frac{\partial^2 \ln L}{\partial \mathrm{vec}(\boldsymbol{B}^s)\partial \mathrm{vec}(\boldsymbol{B}^s)'}\right)^{-1} \left(\frac{\partial \ln L}{\partial \mathrm{vec}(\boldsymbol{B}^s)}\right)$$

Die endgültige Lösung ist gefunden, wenn in zwei aufeinanderfolgenden Iterationsschritten die Minimierungsfunktion (7) nicht weiter abnimmt bzw. der Vektor der Regressionskoef-

---

5 Eine auch für Nichtmathematiker verständliche Erklärung der Logik des Newton-Raphson-Algorithmus findet man **in *Maier u.*** *Weiss* (1990: 84ff.). Meine Darstellung orientiert sich an den Beschreibungen des Algorithmus bei der SAS-Prozedur CATMOD (vgl. *SAS*, 1985: 212f.) und beim Programm KALOS (vgl. ***Röding*** *u.a.* 1985, 5.2-1 ff.).

fizienten sich nicht mehr ändert, da der Vektor der ersten Ableitungen nur Nullen enthält.[6] Als Startwert wird für die Iteration ein Vektor gewählt, bei dem alle Regressionskoeffizienten den Wert Null haben.

Multipliziert man den Wert der Minimierungsfunktion mit 2, ergeben sich die Ausgangsgrößen für die Chiquadrat-Differenzentests. Gleichsam als Nebenprodukt führt der Newton-Raphson-Algorithmus außerdem zu den Schätzungen der Standardfehler und Korrelationen der Schätzer. Die Varianz/Kovarianz-Matrix der Schätzer der Regressionskoeffizienten ist nämlich gerade das Negative der Inversen der Matrix der zweiten Ableitungen aus Gleichung (10).

## 2. Erfahrungen bei der Programmierung mit SPSS$^X$-MATRIX

Die SPSS$^X$-Prozedur MATRIX läßt sich als Interpreter einer spezifischen Programmiersprache verstehen, deren Datenelemente Matrizen sind. Da alle Operationen und Funktionen über Matrizen definiert sind, lassen sich mathematische Formeln sehr schnell in Programmanweisungen umsetzen. Die bekannte Kleinstquadratlösung der Parameterschätzung einer linearen Regressionsfunktion ließe sich beispielsweise durch folgende Anweisung realisieren:

```
COMPUTE B=INV(T(X)*X)*T(X)*Y
```

Neben der sequentiellen Abarbeitung solcher COMPUTE-Anweisungen sind Iterationen über "LOOP ... END LOOP"-Schleifen und bedingte Verzweigungen über "DO IF ... END IF"-Blöcke programmierbar. Direkte Sprunganweisungen und Unterprogrammstrukturen, wie sie etwa aus der Programmiersprache BASIC mit den Anweisungen "GO TO" bzw. "GO SUB" bekannt sind, werden nicht zur Verfügung gestellt.

Mit den Befehlen "GET" bzw. "SAVE" können Daten zwischen SPSS$^X$-Systemdateien und der Prozedur MATRIX ausgetauscht werden. Dabei ist durch die Angabe der Spezifikation "/MISSING= ..." eine (eingeschränkte) Missing-Data-Behandlung möglich. Die Anweisungen "READ" und "WRITE" ermöglichen den Zugriff auf externe Dateien. Mit der "PRINT'-Anweisung können schließlich Matrizen in die Ausgabedatei eines SPSS$^X$-Jobs geschrieben werden.

---

[6] Tatsächlich wird die Iteration beendet, wenn der relative Funktionsabfall bzw. die ersten Ableitungen kleiner als eine vorgegebene kritische Größe sind.

Theoretisch erlauben diese Programm-Elemente der Prozedur MATRIX eine schnelle Umsetzung des Newton-Raphson-Algorithmus zur Schätzung der Regressionskoeffizienten der multinomialen logistischen Regression. Tatsächlich lassen sich die im letzten Abschnitt vorgestellten Gleichungen auch mühelos in SPSS$^x$-MATRIX-Anweisungen umformen. Als recht hinderlich erweist sich dabei allein eine Restriktion der "COMPUTE"-Anweisung: Auf der linken Seite einer Gleichung muß entweder eine vollständige Matrix oder aber ein einziges Element einer Matrix aufgeführt werden. Sollen dagegen nur die Elemente einer Spalte oder einer Zeile einer Matrix angesprochen werden, müssen dazu für jedes Element eigene COMPUTE-Anweisungen ausgeführt werden.

Trotz des raschen Erfolgs bei der Umsetzung des Algorithmus in ein Matrix-Programm konnte der erste Lösungsversuch noch nicht befriedigen. Es stellte sich nämlich heraus, daß das Programm bereits bei relativ kleinen Stichproben (ab etwa 100 Fällen) unverhältnismäßig viel Zeit zur Berechnung der Lösung benötigte. Der Algorithmus wurde daraufhin so umgeschrieben, daß die Berechnung der Minimierungsfunktion und der ersten und zweiten Ableitungen nicht wie in den Gleichungen (7), (9) und (10) fallweise erfolgt, sondern stattdessen durch Matrizen-Operationen über jeweils die gesamte Stichprobe berechnet wird. Dies führt zu einer erheblichen Beschleunigung der Rechenzeit. Offenbar dauert das fallweise Abarbeiten in einer Schleife erheblich länger als die Ausführung einer Operation über Matrizen.

Bei einem solchen stichprobenweisen Vorgehen muß jedoch darauf geachtet werden, daß die einzelnen Matrizen nicht zu groß werden. Zwar kann die SPSS$^x$-Prozedur MATRIX durchaus Matrizen mit etlichen tausend Elementen bearbeiten. Operationen auf solchen Matrizen können jedoch auch den größten Speicher sprengen. Sind etwa die Elemente jeder Zeile einer Matrix mit einer anderen Zahl zu multiplizieren, kann dies in Matrixschreibweise so formuliert werden, daß die Ursprungsmatrix von links mit einer Diagonalmatrix mit den entsprechenden Faktoren in der Diagonale multipliziert wird. Bei z.B. 3000 Zeilen der Ursprungsmatrix hätte die Diagonal-Matrix entsprechend 3000 mal 3000 Elemente. In solchen Situationen ist es notwendig, vor der Umsetzung von Matrizenoperationen in ein SPSS$^x$-Matrixprogramm die Operationen umzuformen. Statt also etwa von links mit einer Diagonalmatrix zu multiplizieren, ist alternativ auch eine elementweise Multiplikation mit einer Matrix gleicher Größe wie die Ausgangsmatrix möglich, wobei die Zeilen der Multiplikationsmatrix die entsprechenden Faktoren enthalten. Solche Umformungen verlangen jedoch mehr als rudimentäre Kenntnisse des Matrizenkalküls, was den Kreis der Programmieranwender der Prozedur MATRIX vermutlich stark einschränkt.

Neben der Notwendigkeit, in Matrizenschreibweise dargestellte Algorithmen umzuformulieren, um Verarbeitungszeit zu gewinnen oder Speicherplatz einzusparen, traten bei meinem Versuch, die logistische Regression als SPSS^X^-Matrix-Programm zu schreiben, noch einige kleinere Probleme auf, die daher rührten, daß die vorliegende Dokumentation zur Prozedur MATRIX recht knapp gehalten ist und teilweise auch Fehler enthält.[7] So wird etwa in der Kurzbeschreibung der Matrix-Funktion "ALL" behauptet, daß diese Funktion prüft, ob alle Elemente einer Matrix positiv sind. In der detaillierteren Beschreibung der Funktion stellt sich dann heraus, daß diese Funktion tatsächlich prüft, ob alle Elemente einer Matrix ungleich Null sind bzw. mindestens ein Element den Wert Null hat. Ein weiteres Problem trat dadurch auf, daß in der Dokumentation von einer amerikanischen Tastatur ausgegangen wird. Sonderzeichen wie geschweifte Klammern und selbst das Ausrufungszeichen zur Identifikation von Makroschlüsselwörtern mußten auf meiner deutschen Tastatur durch Umlaute dargestellt werden, wobei die "passenden" Umlaute aber zunächst einmal gefunden werden mußten.

Insgesamt traten jedoch keine wirklich schwerwiegenden Probleme bei der Umsetzung des Algorithmus in ein SPSS$^X$-MATRIX-Programm auf.[8] Das Programm wurde daraufhin noch um einige Zusatzfunktionen erweitert, die zusätzliche Koeffizienten berechnen, Signifikanztests durchführen und deskriptive Statistiken ausgeben. Schließlich wurde um das Matrix-Programm ein SPSS$^X$-Makro geschrieben, so daß das Programm wie eine SPSS$^X$-Prozedur aufgerufen werden kann. Als fehlerträchtig zeigte sich hier, daß SPSS$^X$ bei MAKRO- bzw. MATRIX-Anweisungen bisweilen entgegengesetzte Syntaxregeln verwendet. So wird das Ende von Schleifen und bedingten Blöcken in MATRIX-Anweisungen durch "END LOOP" bzw. "END IF" angezeigt, in MAKRO-Befehlen dagegen durch "!DOEND" bzw. "!IFEND". Darüber hinaus scheint es möglicherweise Probleme beim mehrfachen Wechsel zwischen MAKRO- und MATRIX-Anweisungen zum Programmablauf zu geben. Jedenfalls lief mein Programm erst fehlerfrei, nachdem alle "LOOP ... END LOOP"- und "DO IF ... END IF"-Blöcke des MATRIX-Programms innerhalb eines einzigen "!IF ... !IFEND "-Blocks der MAKRO-Umgebung untergebracht waren.

Als Ergebnis des Programmierversuchs ist schließlich ein SPSS$^X$-Makro entstanden, das wie eine standardmäßige SPSS$^X$-Prozedur mit einem Namen und nachfolgenden Spezifikationen aufgerufen werden kann (siehe Anhang 1). Allerdings ist die Ausgabe dieser selbsterzeugten Prozedur graphisch nicht so schön wie bei den standardmäßigen SPSS$^X$-Prozeduren. Ursache ist die geringe Mächtigkeit des "PRINT"-Befehls. So können mit einem

---

7 Ich konnte mich bei meinen Versuchen nur auf den "SPSS Reference Guide" stützen. Möglicherweise enthält der "SPSS Advanced Statistics User's Guide" eine umfassendere Dokumentation der SPSS$^X$-Prozedur MATRIX.

8 Für Interessierte sind die Programmanweisungen zur Parameterschätzung im Anhang 2 wiedergegeben.

"PRINT'-Befehl nicht mehrere Matrizen angesprochen werden. Außerdem kann nur ein einziges Format für alle Elemente der auszugebenen Matrix spezifiziert werden.

Die Größe des Makros wuchs im Laufe der Zeit auf über 500 Zeilen an. Da in der Beschreibung zur $SPSS^x$-Prozedur MATRIX und zum MAKRO-Interpreter keine Angaben über die maximale Anzahl der Anweisungen innerhalb eines Matrixprogramms bzw. eines Makros gegeben werden, ist unklar, wie groß ein solches Programm maximal sein kann. Auch der Matrix-Befehl "DISPLAY STATUS" war hier nicht hilfreich, zumal mir keine Informationen über die Interpretation der Ausgabe dieses Befehls vorlagen. Unzutreffende Kommentare bei Syntaxfehlern[9] am Ende meiner Programmierversuche weisen darauf hin, daß möglicherweise die Kapazitätsgrenze erreicht wurde. Ähnlich dürfte auch das Phänomen zu deuten sein, wonach unerklärliche Fehlerabbrüche nicht mehr auftraten, wenn vor dem Einbinden und Abarbeiten des Makros mit dem "EXEC'-Befehl Datenmodifikationen ohne nachfolgenden Prozeduraufruf ausgeführt wurden. Überraschenderweise stellte sich auch heraus, daß bei wiederholtem Aufrufen des Makros innerhalb eines $SPSS^x$-Jobs ab dem dritten Aufruf das Programm mit einer (nicht zutreffenden) Syntaxfehlermeldung abbrach. Eine bessere Dokumentation oder gar eine vom Benutzer beeinflußbare dynamische Alloziierung von Programmspeicherplatz wäre bei der praktischen Arbeit sicherlich sehr hilfreich.

Zur Prüfung der Brauchbarkeit des Programms in (typischen) sozialwissenschaftlichen Analysen wurden zunächst mit den Daten des kumulierten ALLBUS einige logistische Regressionsmodelle geschätzt. Bei diesen Tests traten zunächst neue Probleme auf. Diese betreffen Funktionen der $SPSS^x$-Prozedur MATRIX. So können einige sehr große Funktionswerte bei der Berechnung der kumulativen Chiquadratverteilung zu Fehlerabbrüchen führen. Weiter zeigte sich, daß die Funktionsberechnung der Determinante einer Matrix gelegentlich mit der Fehlermeldung abbrach, daß die Determinante nicht darstellbar sei. Bei dem Versuch, diesen Fehler dadurch zu umgehen, daß statt der Determinante der Rank einer Matrix berechnet wurde, stellte sich heraus, daß die Matrixfunktion "RANK" offenbar auch nicht immer den tatsächlichen Rank einer Matrix korrekt berechnete. Bei meinem Programm zur logistischen Regression konnten diese Probleme schließlich über die Analyse der Eigenwerte einer Matrix umgangen werden. So führten diese "Kinderkrankheiten" der $SPSS^x$-Prozedur MATRIX in meinem Anwendungsfall nicht zu falschen Ergebnissen. Jedenfalls waren alle Berechnungen mit der selbstgeschriebenen Prozedur stets identisch mit denen, die zur Kontrolle auch mit der SAS-Prozedur CATMOD berechnet wurden.

---

9 Recht bizarr war beispielweise eine Fehlermeldung, bei der statt einer Beschreibung des Fehlers der "HELP"-Text der $SPSS^x$-Prozedur MANOVA ausgegeben wurde.

Bei den Testläufen stellte sich heraus, daß auch große Datensätze verarbeitet werden konnten. So waren durch Heraufsetzen der Obergrenze des virtuellen Arbeitsspeichers auch die Datenmengen des kumulierten ALLBUS mit der $SPSS^X$-Prozedur MATRIX zu analysieren. Selbst ein Modell mit mehr als 7200 Fällen, 29 erklärenden Variablen und einer abhängigen Variable mit vier Kategorien konnte das Programm bewältigen. Es benötigte allerdings auf dem ZA-Rechner vom Typ IBM 4183 unter CMS 6 relativ viel Zeit zur Berechnung der Lösung.[10] Insgesamt ist das Ergebniss der Programmierung positiv zu bewerten: Nach der Beseitigung der Anfangsprobleme steht nun ein prozedurartiges $SPSS^X$-Makro zur Verfügung, das allgemein eingesetzt werden kann.

## 3. Eine realistische Anwendung: Die Prüfung einer Hypothese der Theorie von *Fishbein und Ajzen*

Eine der prominentesten Handlungstheorien der Sozialpsychologie ist die "Theorie bedachter Handlungen" (theory of reasoned action) von *Fishbein und Ajzen* (1975; *Ajzen und Fishbein,* 1980; *Ajzen,* 1988). Nach dieser Theorie sind Handlungen als Realisationen von Handlungsabsichten aufzufassen, welche wiederum durch die Einstellung zum Verhalten und die subjektive Norm bezüglich des Verhaltens erklärt werden. Andere Variablen können als externe Faktoren allein indirekt über die Beinflussung der Einstellung, der subjektiven Norm oder des relativen Gewichts von Einstellung respektive Norm auf die Handlungsabsicht und deren Ausführung wirken.

Mit den Daten, die im Rahmen der Begleituntersuchung zur Volkszählung 1987 erhoben wurden (vgl. *Scheuch u.a.,* 1989), soll im folgenden die Hypothese geprüft werden, daß allein die Verhaltensabsicht einen signifikanten Einfluß auf das Teilnahmeverhalten bei der Volkszählung 1987 hatte. Insbesondere sollte nach der Theorie bedachter Handlungen die Einstellung zur Volkszählung keinen eigenständigen Einfluß neben der Verhaltensabsicht haben. In der Begleituntersuchung zur Volkszählung 1987 wurde kurz vor dem Beginn der Volkszählung u.a. die Verhaltensabsicht erhoben. Ein gutes halbes Jahr später gaben die Befragten Auskunft über ihr tatsächlich realisiertes Verhalten. Bei der Handlungsabsicht wie beim späteren Verhaltensbericht konnten die Befragten zwischen fünf Antwortkategorien wählen: a) alle Fragen so gut wie möglich beantworten, b) einige Fragen nicht beantworten, c) einige Fragen nicht wahrheitsgemäß beantworten, d) den Fragebogen weitgehend nicht wahrheitsgemäß beantworten und e) bestimmt nicht teilnehmen bzw. den Fragebogen nicht beantworten.

---

[10] Eine genaue Angabe über die benötigte Rechenzeit konnte nicht ermittelt werden, da die vom $SPSS^X$-System ausgedruckten CPU-Zeiten bei einigen Jobs offensichtlich nicht zutrafen.

In der folgenden Analyse wird die Wahl der ersten Antwortkategorie als *Kooperation,* die der letzten als *offener Boykott* und eine Antwort in einer der drei mittleren Kategorien als *verdeckter Boykott* bezeichnet. Die Zusammenfassung der mittleren Kategorien läßt sich damit rechtfertigen, daß bei diesen Handlungsalternativen die Aufforderung der Zählstellen zur Teilnahme nicht vollständig ignoriert wird. Man kann die Ansicht vertreten, daß die drei Kategorien eine ordinale Variable definieren und entsprechend das logistische Modell für ordinale Daten anwenden, das *Ludwig-Mayerhofer* in seinem Beitrag in diesem Heft vorstellt. Nach den Ergebnissen von anderen Analysen dieser Daten *(Kühnel u. Scheuch,* 1990) scheint die Auffassung einer nominalskalierten Variablen allerdings angemessener zu sein.

Die affektiv-emotionale Einstellung zur Volkszählung wurde in der Begleituntersuchung durch die Frage nach der generellen Haltung zur Zählung erfaßt, wobei sieben Antwortmöglichkeiten von *"sehr ablehnend"* (1) bis *"sehr zustimmend"* (7) vorgegeben waren. Nach der Theorie von *Fishbein und Ajzen* sollte die vor der Volkszählung erhobene Einstellung nach der Kontrolle der Verhaltensabsicht keinen Effekt auf das (berichtete) Teilnahmeverhalten haben. Zur Prüfung dieser Hypothese habe ich mit meinem Matrixprogramm ein logistisches Regressionsmodell analysiert, in dem das nach der Zählung berichtete Teilnahmeverhalten durch die vor der Zählung erhobene Einstellung zur Volkszählung und die Teilnahmeabsicht erklärt wird. Da die Verhaltensabsicht wie das berichtete Antwortverhalten als eine trichotome nominalskalierte Variable aufgefaßt wird, ist diese erklärende Variable in zwei dichotome Dummy-Variablen aufgespaltet worden, die den Wert Eins annehmen, wenn ein Befragter als Verhaltensabsicht offenen Boykott (OFFEN) bzw. verdeckten Boykott (VERDECKT) angibt. Befragte, die sich kooperativ verhalten wollen, haben bei beiden Variablen den Wert Null. Zusammen mit der Regressionskonstante hat das Regressionsmodell somit insgesamt vier Prädiktoren. Dieses Modell wird durch die SPSS$^x$-Anweisung:

```
MLR VAR=VZTEILN OFFEN VERDECKT VZEINST1 /RKAT=0
```

berechnet, wobei "VZTEILN" der Variablenname für das berichtete Antwortverhalten ist und "OFFEN", "VERDECKT" und "VZEINST1" die drei erklärenden Variablen sind.[11]

Tabelle 1 zeigt eine Zusammenfassung der vom Programm ausgedruckten Koeffizienten. Da die abhängige Variable drei Kategorien aufweist, sind jedem Prädiktor zwei Regressionskoeffizienten zugeordnet, so daß das Modell insgesamt 8 Parameter aufweist. Der

[11] Eine vollständige Beschreibung des Aufrufs des Matrixprogramms befindet sich in Anhang 1 dieses Beitrags.

Tabelle ist zu entnehmen, daß sich die Devianz, d.h. -2 mal der Wert der logarithmierten Likelihood-Funktion (5) bei 1156 gültigen Fällen von 2539,99 auf 610,04 verringert. Eine ganz erhebliche Devianzreduktion erhält man jedoch auch schon in einem Modell, das nur die Regressionskonstante enthält. Bezogen auf dieses Konstantenmodell verringert sich die Devianz nur um 134,74. Bei sechs Freiheitsgraden (den sechs Regressionskoeffizienten der beiden Dummy-Variablen und der Einstellung zur Volkszählung) ist dieser chiquadratverteilte Wert signifikant von Null verschieden. Anteilsmäßig beträgt die Devianzreduktion gegenüber dem Konstantenmodell 18,1 %. Dieser Prozentwert kann als Analogon zum Determinationskoeffizienten in einem linearen Regressionsmodell aufgefaßt werden.

Da zu jedem Prädiktor zwei Regressionskoeffizienten gehören, ergibt sich für den Signifikanztest der Prädiktoren eine Chiquadratverteilung mit zwei Freiheitsgraden. Die Nullhypothese postuliert, daß beide Koeffizienten Null sind. Obwohl die Werte der Regressionskoeffizienten von der gewählten Referenzkategorie abhängen, führt der Test bei beliebigen Referenzkategorien zum gleichen Ergebnis. Tabelle 1 ist zu entnehmen, daß alle Prädiktoren einen signifikanten Einfluß auf die abhängige Variable haben. Inhaltlich bedeutet dies insbesondere, daß auch die Einstellung zur Volkszählung neben der Verhaltensabsicht einen eigenen direkten Einfluß auf das (berichtete) Teilnahmeverhalten hat, die Hypothese aus der Theorie von *Fishbein und Ajzen* demnach falsch zu sein scheint.

Der Blick auf die Regressionskoeffizienten zeigt, in welche Richtung die erklärenden Variablen wirken. Wird die Kategorie *"Kooperation"* als Referenzkategorie gewählt, so sieht man am negativen Vorzeichen der Regressionskoeffizienten, daß sich bei einem höheren Wert auf der Variable "VZ-Einstellung" das Wahrscheinlichkeitsverhältnis zuungunsten der Antwortkategorien *"verdeckter Boykott"* (ß=-0.421) bzw. *"offener Boykott"* (ß=-0.477) entwickelt. Aus den T-Werten, die sich aus den Quotienten der Regressionskoeffizienten, geteilt durch die zugehörigen Standardfehler, ergeben, wird ersichtlich, daß beide Koeffizienten auf dem 5%-Niveau signifikant von Null verschieden sind. Die Exponentation der Regressionskoeffizienten wird in der Tabelle 1 als Effekt bezeichnet (vgl. *Long,* 1987). Die Werte geben an, um welchen Faktor sich das Wahrscheinlichkeitsverhältnis ändert, wenn die erklärende Variable um +1 Einheit zunimmt. Zum Vergleich der relativen Einflüsse der Variablen geben die standardisierten Effekte an, wie sich die Wahrscheinlichkeitsverhältnisse bei einer Änderung eines Prädiktor um +1 Standardabweichung ändern. Bei Werten kleiner Eins ist es sinnvoll, den Kehrwert anzugeben.

Der Tabelle ist zu entnehmen, daß die Chancen der *"Kooperation"* gegenüber dem *"verdeckten Boykott"* um etwa das Eineinhalbfache (1.52) steigen, wenn die Einstellung um +1 Einheit zunimmt. Gegenüber dem *"offenen Boykott"* steigt die Chance der *"Kooperation"* sogar noch etwas stärker (um den Faktor 1,61). Inhaltlich bedeutet dies, daß mit zuneh-

**Tabelle 1:** Logistische Regression des berichteten VZ-Teünahmeverhaltens auf die Teilnahmeabsicht und die Einstellung zur Volkszählung (vor der Zählung)

**Modellanpassung**:

| Modell | Parameter zahl | Devianz |
|---|---|---|
| Nullmodell | 0 | 2539.99 |
| Konstantenmodell | 2 | 744.78 |
| Endmodell | 8 | 610.04 |

Devianzreduktion: 18.1 % (Chiquadrat: 134.74, df: 6, P < 0.0005)

Fallzahl: 1156

**Signifikanztest der Prädiktoren:**

| Variable | Chiquadrat | df | P |
|---|---|---|---|
| Offen | 19.069 | 2 | <.0005 |
| Verdeckt | 21.096 | 2 | <.0005 |
| VZ-Einstellung | 29.830 | 2 | <.0005 |
| (Konstante) | 34.041 | 2 | <.0005 |

**Regressionskoeffizienten:**

*a) Verdeckter Boykott zu Kooperation*

| Prädiktor | Beta | Std. Fehler | T-Wert | Effekt | Std.Eff. |
|---|---|---|---|---|---|
| Offen | 1.589 | .505 | 3.15 | 4.90 | 1.31 |
| Verdeckt | 1.219 | .293 | 4.16 | 3.38 | 1.78 |
| VZ-Einstellung | -0.421 | .082 | -5.15 | 1/1.52 | 1/2.31 |
| (Konstante) | -1.798 | .380 | -4.74 | 1/6.04 | |

*b) Offener Boykott zu Kooperation*

| Prädiktor | Beta | Std.Fehler | T-Wert | Effekt | Std.Eff. |
|---|---|---|---|---|---|
| Offen | 4.004 | 1.209 | 3.31 | 54.83 | 1.99 |
| Verdeckt | 2.212 | 1.0 90 | 2.03 | 9.13 | 2.84 |
| VZ-Einstellung | -0.477 | .231 | -2.06 | 1/1.61 | 1/2.58 |
| (Konstante) | -4.548 | 1.276 | -3.56 | 1/94.40 | |

*c) Verdeckter Boykott zu offenem Boykott*

| Prädiktor | Beta | Std.Fehler | T-Wert | Effekt | Std.Eff. |
|---|---|---|---|---|---|
| Offen | -2.415 | 1.266 | -1.91 | 1/11.19 | 1/1.51 |
| Verdeckt | -0.993 | 1.123 | -0.88 | 1/2.70 | 1/1.60 |
| VZ-Einstellung | 0.056 | .242 | 0.23 | 1.06 | 1.12 |
| (Konstante) | 2.749 | 1.319 | 2.08 | 15.63 | |

mender positiver Einstellung zur Volkszählung die Wahrscheinlichkeit der *"Kooperation"* gegenüber dem *"verdeckten Boykott"* selbst dann steigt, wenn die Teilnahmeabsicht als Kontrollvariable im Regressionsmodell berücksichtigt wird.

Die unterste Teiltabelle zeigt den Einfluß der Prädiktoren, wenn die Kategorie *"offener Boykott"* als Referenzkategorie gewählt wird. Entsprechend Gleichung (3) ergeben sich die Regressionskoeffizienten aus der Differenz der beiden anderen Tabellen. Zur Berechnung der Standardfehler und der T-Werte benötigt man die Kovarianzen bzw. Korrelationen der Schätzer. Das Matrixprogramm zur logistischen Regression berechnet durch die Angabe der Spezifikation "/RKAT=0" Regressionskoeffizienten, Standardfehler, T-Werte, Wahrscheinlichkeiten und unstandardisierte und standardisierte Effekte für alle möglichen Referenzkategorien der abhängigen Variable.

Bei der Erläuterung der Koeffizienten aus Tabelle 1 wurde bereits darauf hingewiesen, daß die Ergebnisse dieser Analyse als Falsifikation der Hypothese von ***Fishbein und Ajzen*** aufgefaßt werden können, nach der allein die Verhaltensabsicht die Ausführung einer Handlung beeinflußt. Diese Bewertung kann jedoch etwas vorschnell sein. So weisen ***Fishbein und Ajzen*** (1975:370) selber darauf hin, daß diese Beziehung nur gilt, wenn zwischen der Messung der Verhaltensabsicht und der geplanten Verhaltensausführung keine Ereignisse eintreten, die die Verhaltensabsichten verändern. Tatsächlich haben wir in der Begleituntersuchung zur Volkszählung feststellen können, daß bei Beginn der Zählung eine deutliche Änderung der Einstellung zur Zählung zugunsten einer positiveren Bewertung festzustellen war (vgl. ***Kühnel,*** 1988:36; ***Scheuch*** *u.a.,* 1989:33). Wenn diese Veränderungen die Verhaltensabsichten beeinflußt haben, ist die erhobene Verhaltensabsicht **vor** der Zählung nur ein unvollkommener Schätzer der nach der Theorie letztlich relevanten Verhaltensabsicht **bei** der Zählung. Da andererseits die Einstellung vor der Zählung mit der späteren Einstellung während der Zählung korreliert ist, kann die Einstellung vor der Zählung als eine verzerrte Messung der Einstellung bei der Zählung aufgefaßt werden. Da theoriegemäß die Einstellung bei der Zählung die Verhaltensabsicht bei der Zählung beeinflussen kann, ist bei dieser Interpretation durchaus damit zu rechnen, daß die vor der Zählung erhobene Verhaltensabsicht und die Einstellung zur Zählung beide das spätere Teilnahmeverhalten beeinflussen.

Zur Überprüfung dieser Vermutung habe ich in einem zweiten Analyseschritt die nach der Volkszählung erhobene Einstellung zur Zählung als weiteren Prädiktor zur Erklärung des berichteten Antwortverhaltens herangezogen, da aufgrund der genannten früheren Analysen ***(Kühnel,*** 1988; ***Scheuch u.a.***, 1989) diese Messung als relativ reliabler Indikator der Einstellung zur Volkszählung während der Zählung aufgefaßt werden kann. Die Ergebnisse dieser zweiten Regressionsschätzung sind in Tabelle 2 festgehalten.

**Tabelle 2:** Logistische Regression des berichteten VZ-Teilnahmeverhaltens auf die Teilnahmeabsicht und die Einstellung zur Volkszählung (vor und nach der Zählung)

**Modellanpassung:**

| Modell | Parameterzahl | Devianz |
|---|---|---|
| Nullmodell | 0 | 2533.40 |
| Konstantenmodell | 2 | 744.26 |
| Endmodell | 10 | 552.16 |

Devianzreduktion: 25.8 % (Chiquadrat: 192.10, df: 8, P < 0.0005)

Fallzahl: 1153

**Signifikanztest der Prädiktoren:**

| Variable | Chiquadrat | df | P |
|---|---|---|---|
| Offen | 14.549 | 2 | .001 |
| Verdeckt | 17.510 | 2 | <.0005 |
| Einstellung vor VZ | 4.501 | 2 | .105 |
| Einstellung nach VZ | 47.877 | 2 | <.0005 |
| (Konstante) | 5.622 | 2 | .060 |

**Regressionskoeffizienten:**

***a) Verdeckter Boykott zu Kooperation***

| Prädiktor | Beta | Std.Fehler | T-Wert | Effekt | Std.Eff. |
|---|---|---|---|---|---|
| Offen | 1.523 | .548 | 2.78 | 4.59 | 1.30 |
| Verdeckt | 1.133 | .300 | 3.78 | 3.10 | 1.71 |
| Einst. vor VZ | -0.190 | .090 | -2.12 | 1/1.21 | 1/1.46 |
| Einst. nach VZ | -0.485 | .079 | -6.14 | 1/1.62 | 1/2.40 |
| (Konstante) | -0.530 | .425 | -1.25 | 1/1.70 | |

***b) Offener Boykott zu Kooperation***

| Prädiktor | Beta | Std.Fehler | T-Wert | Effekt | Std.Eff. |
|---|---|---|---|---|---|
| Offen | 3.870 | 1.278 | 3.03 | 47.97 | 1.94 |
| Verdeckt | 2.131 | 1.099 | 1.94 | 8.43 | 2.74 |
| Einst. vor VZ | -0.054 | .243 | -0.22 | 1/1.06 | 1/1.11 |
| Einst. nach VZ | -0.840 | .224 | -3.75 | 1/2.32 | 1/4.55 |
| (Konstante) | -2.822 | 1.348 | -2.09 | 1/16.81 | |

***c) Verdeckter Boykott zu offenem Boykott***

| Prädiktor | Beta | Std.Fehler | T-Wert | Effekt | Std.Eff. |
|---|---|---|---|---|---|
| Offen | -2.347 | 1.306 | -1.80 | 1/10.46 | 1/1.50 |
| Verdeckt | -0.999 | 1.128 | -0.89 | 1/2.71 | 1/1.60 |
| Einst. vor VZ | -0.136 | .253 | -0.54 | 1/1.15 | 1/1.31 |
| Einst. nach VZ | 0.355 | .230 | -1.54 | 1.43 | 1.90 |
| (Konstante) | 2.291 | 1.386 | 1.65 | 9.89 | |

Tatsächlich ergibt diese Analyse, daß nun die Einstellung vor der Zählung keinen signifikanten Einfluß auf das berichtete Teilnahmeverhalten hat. Bei zwei Freiheitsgraden ergibt sich für diesen Prädiktor ein Chiquadratwert von 4,5 (P: 0,105). Betrachtet man jedoch die einzelnen Regressionskoeffizienten, so bleibt doch ein leichter Effekt der Einstellung vor der Zählung auf das Wahrscheinlichkeitsverhältnis von *"verdecktem Boykott"* zu *"Kooperation"* sichtbar, der auch auf dem 5%-Niveau signifikant ist (T-Wert: 2.12).

Insgesamt ergibt sich somit ein Ergebnis, das sowohl zugunsten als auch zuungunsten der Theorie von ***Fishbein und Ajzen*** interpretiert werden kann. Aus einer Perspektive, nach der im Sinne von ***Lakatos*** (1982) Theorien erst aufgegeben werden sollen, wenn adäquatere Theorien formuliert sind, läßt sich das Gesamtergebnis m.E. eher als eine Bestätigung der Theorie von ***Fishbein und Ajzen*** auffassen.[12]

## 4. Resümee

In diesem Beitrag habe ich an einem Beispiel untersucht, ob es möglich ist, mit der Prozedur MATRIX Analysen durchzuführen, für die im Rahmen von SPSS$^X$ keine Standardprozedur zur Verfügung gestellt wird. Zumindest für das Anwendungsbeispiel läßt sich diese Frage positiv beantworten. Nach anfänglichen Schwierigkeiten war es möglich, ein SPSS$^X$-Makro zu programmieren, mit dem sich Daten der in der quantitativen Sozialforschung üblichen Größenordnung wie mit einer standardmäßigen SPSS$^x$-Prozedur analysieren lassen. Die Einbindung der Prozedur MATRIX in das SPSS$^x$-System kann daher als eine erhebliche Verbesserung dieses Programmpaketes betrachtet werden.

Diese generelle Beurteilung ist allerdings zu relativieren. Zum einen bin ich bereits in der ersten Anwendung der Prozedur MATRIX auf offensichtliche Programmierfehler gestoßen. Es ist zu hoffen, daß solche Fehler schnell gefunden und in den nächsten Versionen behoben werden. Grundsätzlich vermeiden lassen sich bei einem großen Programmsystem Programmierfehler allerdings kaum. Wünschenswert wäre auch eine umfassendere Dokumentation, die insbesondere auf die Grenzen des Programms eingeht.

Vergleicht man die SPSS$^x$-Prozedur MATRIX mit Konkurrenzprodukten, wie etwa der PROC MATRIX bzw. dem IML-System von SAS, oder mit dem PC-System GAUSS, so erreicht der Matrixinterpreter in SPSS$^x$ noch nicht deren Leistungsumfang. Dies liegt auch daran, daß in die SPSS$^x$-Prozedur MATRIX nicht alle Funktionen implementiert wurden,

---

12 Bei einer inhalüich orientierten Arbeit wäre an dieser Stelle eine Einbeziehung weiterer erklärender Variablen notwendig. Eine umfassende Analyse würde jedoch den Rahmen dieses eher methodischen Beitrags sprengen.

die in SPSS$^X$ bei der Aufbereitung der Daten zur Verfügung stehen. Eine ernsthafte Einschränkung der Anwendungsmöglichkeiten sehe ich hierin allerdings nicht.

Anwendungsgrenzen dürften eher darin bestehen, daß die Beherrschung des Matrizenkalküls nicht zu den Techniken gehört, die üblicherweise von empirisch arbeitenden Sozialwissenschaftlern verlangt werden. Soll das mit der Prozedur MATRIX bereitgestellte Instrument über den Kreis der Methodenexperten hinaus genutzt werden, wird es notwendig sein, von Spezialisten geschriebene Matrixprogramme einem breiteren Publikum zur Verfügung zu stellen.

## Literatur:


***Ajzen, I.*** (1988), Attitudes, Personality, and Behavior. Stony Stratford: Open University.

***Ajzen, I. u. M. Fishbein*** (1980), Understanding Attitudes and Predicting Social Behavior. Englewood Cliffs: Prentice-Hall.

***Fishbein, M. u. I. Ajzen*** (1975), Belief, Attitüde, Intention and Behavior: An Introduction to Theory and Research. Reading, Mass.: Addison-Wesley.

***Ludwig-Mayerhofer, W.*** (1990), "Multivariate Logit-Modelle für ordinalskalierte abhängige Variablen. ZA-Information 27 (in diesem Heft).

***Maier, G. u. P. Weiss*** (1990), Modelle diskreter Entscheidungen. Theorie und Anwendung in den Sozial- und Wirtschaftswissenschaften. Wien: Springer.

***Kühnel, SM.*** (1988), "Full-Quasi-Likelihood-Schätzung: Die Ausnutzung aller verfügbaren Informationen bei der Schätzung von linearen Strukturgleichungsmodellen." ZA-Information 23:24-46

***Kühnel, SM., W. Jagodzinski u. M. Terwey*** (1989), "Teilnehmen oder Boykottieren: Ein Anwendungsbeispiel der binären logistischen Regression mit SPSS$^X$." ZA-Information 25: 44-75.

***Kühnel, SM. u. E.K. Scheuch*** (1990), Der Einfluß der Strafandrohung auf das Teilnahmeverhalten bei der Volkszählung. Ergebnisse der Kölner Begleituntersuchung zur Volkszählung. Köln: Unveröffentlichtes Vortragsmanuskript.

***Lakatos, I.*** (1982), "Falsifikation und die Methodologie wissenschaftlicher Forschungsprogramme." S. 7-107 in: /. ***Lakatos,*** Philosophische Schriften, herausgegeben von ***J. Worral und G. Currie,*** Band 1: Die Methodologie der wissenschaftlichen Forschungsprogramme. Braunschweig: Vieweg.

***Long, J.S.*** (1987), "A Graphical Method for the Interpretation of Multinomial Logit Analysis." Sociological Methods and Research, 15:420-446.

***Röding, M., U. Küsters u. G. Arminger*** (1985), KALOS Version 1.0. Ein interaktives Programmsystem zur Analyse kategorialer Logitmodelle. Programmbeschreibung. Köln: ZA.

***SAS Institute Inc.*** (1985), SAS User's Guide: Statistics, Version 5 Edition. Cary. SAS Institute Inc.

***Scheuch, E., L. Graf u. S. Kühnel*** (1989), Volkszählung, Volkszählungsprotest und Bürgerverhalten. Ergebnisse der Begleituntersuchung zur Volkszählung 1987. Stuttgart: Metzler-Poeschel.

***SPSS Inc.*** (1990), SPSS Reference Guide. Chicago: SPSS Inc.

***Urban, D.*** (1990), "Multinomiale LOGIT-Modelle zur Bestimmung der Abhängigkeitsstruktur qualitativer Variablen mit mehr als zwei Ausprägungen. ZA-Information 26: 36-61.

## ANHANG 1: Der Aufruf des SPSS$^x$-Makro "MLR"

Das für diesen Beitrag entwickelte Matrixprogramm zur ML-Schätzung multinomialer logistischer Regressionsmodelle kann als SPSS$^x$-Makro wie eine Prozedur aufgerufen werden. Nach dem Makronamen "MLR", der wie jedes SPSS$^x$-Kommando auf der ersten Spalte einer Zeile beginnen muß, folgen entsprechend der üblichen Syntax von SPSS$^X$ eine oder mehrere Spezifikationen, die jeweils durch einen Schrägstrich getrennt werden müssen. Die Spezifikationen können auf derselben Zeile oder in Folgezeilen stehen. Sie dürfen bei Fortsetzungen über mehrere Zeilen nicht in der ersten Spalte beginnen. Im Unterschied zu Standardprozeduren können die Unterbefehle nicht auf drei Zeichen abgekürzt werden.

Die vollständige Syntax des Makros sieht folgendermaßen aus:

```
MLR  VAR = variablenliste
     /RKAT = ymax
     /NOCONST
     /COR
     /DES
     /NIT = 25     /EPS1 = 0.0000001     /EPS2 = 0.0001
     /NCASE = 20000     /PRESORT
```

Notwendig ist nach dem Makroaufruf einzig und allein die Angabe einer Variablenliste nach dem Schlüsselwort "VAR=". Für die Variablenliste gelten die üblichen SPSS$^x$-Konventionen. Variablennamen können durch Leerstelle oder Komma getrennt werden. Das Schlüsselwort "TO" kann verwendet werden. Gegenüber anderen Prozeduren besteht die Einschränkung, daß ein Variablenname nur einmal in der Variablenliste aufgeführt werden darf. Anderenfalls bricht das Programm mit einer SPSS$^x$-MATRIX-Fehlermeldung ab.

Das Matrixprogramm verlangt, daß die erste der aufgeführten Variablen die nominalskalierte abhängige Variable ist und alle folgenden Variablen die erklärenden Variablen sind. Zulässige Werte für die abhängige Variable sind nur ganze Zahlen, wobei der kleinste Wert "1" sein muß. Das Programm verlangt weiter, daß alle Kategorien von der kleinsten bis zur größten vorgefundenen ganzen Zahl Besetzungszahlen größer Null aufweisen. Gegebenenfalls muß die abhängige Variable vor dem Aufruf der Prozedur rekodiert werden. Das Makro prüft diese Bedingung ab und gibt gegebenenfalls eine entsprechende Fehlermeldung aus. Da grundsätzlich jeder Fall von der Analyse ausgeschlossen wird, wenn er auf mindestens einer der in der Variablenliste aufgeführten Variablen einen ungültigen Wert aufweist, kann ein Fehlerabbruch auch dann erfolgen, wenn eine univariate Häufigkeitsauszählung der abhängigen Variable in jeder Kategorie Besetzungen zeigt.

Das Modell der logistischen Regression unterstellt, daß die erklärenden Variablen metrisch sind. Statt einer kategorialen erklärenden Variable sind daher in der Variablenliste des Makros gegebenenfalls dichotome Dummy-Variablen aufzuführen. Dabei ist darauf zu achten, daß bei m Kategorien der erklärenden Variable maximal (m-1) Dummy-Variablen aufgeführt werden. Das Makro druckt eine Warnung aus, wenn der kleinste Eigenwert der Rohproduktmomentmatrix kleiner als 10* ist. Wenn die Matrix der zweiten Ableitungen während der Iterationen singulär wird, wird der Minimierungsprozeß mit einer entsprechenden Meldung abgebrochen.

Hinter dem Schlüsselwort "/RKAT =" kann die Referenzkategorie für die Koeffizienten der abhängigen Variable spezifiziert werden. Wird keine Referenzkategorie aufgeführt, wird als Voreinstellung der größte Wert der abhängigen Variable gewählt (oben durch "ymax" symbolisiert). Die Angabe "RKAT = 0" bewirkt, daß nacheinander jede Kategorie der abhängigen Variable als Referenzkategorie betrachtet wird und so alle theoretisch möglichen Regressionskoeffizienten ausgegeben werden.

Ergänzend zu den in der Variablenliste aufgeführten erklärenden Variablen wird standardmäßig eine Regressionskonstante in die Regressionsgleichung aufgenommen. Durch Angabe des Parameters "/NOCONST" kann ein Modell ohne Regressionskonstante geschätzt werden.

Neben den Regressionskoeffizienten und einigen daraus abgeleiteten Kenngrößen werden standardmäßig Signifikanztests der einzelnen Prädiktoren berechnet. Außerdem werden Chiquadrat-Differenzentests und Pseudo-$R^2$-Werte berechnet. Nullmodell ist dabei zum einen das Ausgangsmodell der Iteration, bei dem alle Regressionskoeffizienten den Wert Null haben, und zum anderen ein Modell, in dem nur die Regressionskonstante enthalten ist.[1] Zur Kontrolle des Minimierungsverlaufs wird außerdem ein Protokoll der Minimierungsschritte ausgedruckt.

Zusätzliche Ausgaben können über weitere Schlüsselwörter angefordert werden. Mit der Spezifikation "/COR" können die Korrelationen der Schätzungen ausgegeben werden. Durch Angabe des Parameters "/DES" werden deskriptive Statistiken ausgegeben. Diese umfassen eine Häufigkeitstabelle der abhängigen Variable und die Mittelwerte, Standardabweichungen und Korrelationen der erklärenden Variablen.

Drei weitere Schlüsselwörter beeinflussen den Schätzalgorithmus. Durch die Angabe von "/NIT = k" wird die maximale Anzahl von Iterationen auf k festgelegt. Als Voreinstellung sind 25 Iterationen vorgesehen. Wenn der Algorithmus nicht in der vorgegebenen Maximalzahl von Iterationen konvergiert, erfolgt eine Fehlermeldung. Die Konvergenz wird über zwei Konvergenzparameter gesteuert. Konvergenz wird zum einen angenommen, wenn der relative Abfall der Minimierungsfunktion zwischen zwei Iterationen kleiner als ein Konvergenzkriterium "EPS1" ist. Der voreingestellte Wert von $10^{-7}$ kann durch den Parameter "/EPS1 = ..." verändert werden. Das Matrixprogramm nimmt weiterhin Konvergenz an, wenn das betragsmäßig größte Element des Vektors der ersten Ableitungen der Regressionskoeffizienten kleiner als ein kritischer Wert "EPS2" ist Die Voreinstellung für diesen Wert ist $10^{-4}$ Wie das Kriterium "EPS 1" kann auch dieser Wert überschrieben werden.

Zur Berechnung von $d_{it}$ (vgl. Gleichung 7 und 9) für jede Kategorie der abhängigen Variable muß die abhängige Variable einmalig fallweise abgearbeitet werden. Bei mehr als 20 000 Fällen muß mit der Spezifikation "/NCASE = nmax" eine größere Fallzahl angegeben werden.[2] Sind die Daten nach den Werten der abhängigen Variable aufsteigend sortiert, kann das fallweise Abarbeiten entfallen. In diesem Fall kann durch Angabe der Spezifikation "/PRESORT" die Rechenzeit beschleunigt werden.

Aufgrund der Länge des Makros ist es sinnvoll, das Matrixprogramm als externe Datei zu speichern und mit dem $SPSS^x$-Befehl INCLUDE bereitzustellen. Da am Anfang des Programms ein SET-Befehl ausgeführt wird, sollte vor dem INCLUDE-Befehl mit dem Befehl PRESERVE die ursprüngliche $SPSS^x$-Umgebung gesichert werden und nach dem Aufruf des Makros mit dem Befehl RESTORE wieder der Ursprungszustand hergestellt werden. Um den Arbeitsspeicher voll nutzen zu können, ist vor dem INCLUDE-Befehl der Befehl EXEC nützlich.

Insgesamt ergibt sich somit folgende Befehlssequenz zur Berechnung der multinomialen logistischen Regression:

```
PRESERVE
EXEC
INCLUDE FTLE="dateiname der datei mit dem MLR-makro"
MLR VAR=...
RESTORE
```

Bei einem interaktiven Arbeiten mit $SPSS^X$ ist hinter jedem Befehl das Zeichen für den Befehlsabschluß - in der Regel ein Punkt - zu setzen. Treten beim Programmaufruf Fehlermeldungen des MATRIX-Interpreters auf, die auf Programmzeilen hinweisen, so können diese lokalisiert und möglicherweise behoben werden, wenn direkt vor dem Aufruf des Makros der $SPSS^x$-Befehl "SET MPR ON" eingefügt wird.

---

1 Das Ausgangsmodell entspricht der Hypothese, daß die abhängige Variable in der Population gleichverteilt ist Im Konstantenmodell wird zwar die empirische Randverteilung der abhängigen Variable reproduziert, jedoch Unabhängigkeit von den erklärenden Variablen unterstellt.

2 Der Parameter "NCASE" setzt die Anzahl der maximalen Loops im SET-Kommando von $SPSS^X$ hoch.

## ANHANG 2: SPSS$^X$-Anweisungen zur Schätzung der Modellparameter

```
*********** MINIMIERUNG NACH NEWTON/RAPHSON-ALGORITHMUS. ************
COMPUTE FIT=CHI(1,1)                /* STARTWERT DER MIN.-FUNKTION
COMPUTE LOG(1,2)=FIT                 /* AUSGANGSFIT
COMPUTE BETA=MAKE(K,IM1,0)          /* STARTWERTE VON BETA IN 1. ITER.
COMPUTE PROB=MAKE(N,I,(1/I))        /* STARTWERTE DER WAHRSCH.
**** BEGINN MINIMIERUNGSITERATION
PRINT /TIT "*** MINIMIERUNG VON -2 LOG-LIKELIHOOD-FUNKTION"
      /SPA NEW
COMPUTE CONV=0
LOOP ITER=1 TO NITER                    /* MINIMIERUNGSSCHLEIFE
COMPUTE LOG((ITER+1),1)=ITER            /* FESTHALTEN DER ITER.-NUMMER
COMPUTE FOLD=FIT                        /* LETZTER WERT DES FITS
*** BERECHNUNG DER 1. U. 2. ABLEITUNG ***
** A. BERECHNEN DER 1. ABLEITUNG -> DEV1
COMPUTE DEV1=T(X)*(DES(:,1:IM1)-PROB(:,1:IM1))
** MINIMIERUNG ABGESCHLOSSEN, WENN 1. ABLEITUNG NULL
COMPUTE LOG((ITER+1),4)=MMAX(ABS(DEV1))
DO IF (LOG((ITER+1),4) LE EPS2)
PRINT /TITLE="***** KONVERGENZ: 1. ABLEITUNG IST NULL *****"
      /SPACE=5
COMPUTE CONV=1
END IF
** B. BERECHNEN DER INFORMATIONSMATRIX (- 2. ABLEITUNG) -> DEV2
COMPUTE DEV2=MAKE((K*IM1),(K*IM1),0)
LOOP L1=1 TO IM1
LOOP L2=1 TO L1
LOOP L3=1 TO K
DO IF (L1=1)
COMPUTE L4E=L3
ELSE
COMPUTE L4E=K
END IF
LOOP L4=1 TO L4E
DO IF (L1=L2)                             /* WENN PROB(I)=PROB(J)
COMPUTE S=PROB(:,L1)&*(1-PROB(:,L1))      /* S = VARIANZ VON P(I)
ELSE                                      /* WENN PROB(I) <> PROB(J)
COMPUTE S=-1*PROB(:,L1)&*PROB(:,L2)       /* S = KOV<PROB(I),PROB(J)>
END IF
COMPUTE DEV2((L3+K*(L1-1)),(L4+K*(L2-1)))=       /*MULTIPLIKATION MIT*/
        MSUM(X(:,L3)&*S&*X(:,L4))               /* X(K) UND X(L) */
COMPUTE DEV2((L4+K*(L2-1)),(L3+K*(L1-1)))=
        DEV2((L3+K*(L1-1)),(L4+K*(L2-1)))
END LOOP
END LOOP
END LOOP
END LOOP
RELEASE S
*** BERECHNUNG DER NEUEN BETA-WERTE UND DER LIKELIHOODFUNKTION
** A. ABBRUCH, WENN 2. ABLEITUNG NICHT POSITIV DEFINIT
DO IF (ALL(EVAL(DEV2)) = 0)
PRINT ITER /FOR "F2.0" /TITLE="++++ PROBLEM IN ITERATION:"
        /SPACE=5
PRINT /TITLE="     INFORMATIONSMATRIX NICHT POSITIV DEFINIT!"
        /SPACE=1
PRINT /TIT "+++++ PROGRAMM STOPPT MINIMIERUNG!"
COMPUTE ERRMIN=1
BREAK
```

```
** B. KORREKTUMATRIX, WENN 2. ABLEITUNG POSITIV DEFINIT
                GRAD = -INV(DEV2)*DEV1
ELSE
COMPUTE DEV1=RESHAPE(T(DEV1),(K*IM1),1)
COMPUTE GRAD=INV(DEV2)*DEV1
COMPUTE GRAD=T(RESHAPE(GRAD,IM1,K))
END IF
** C. NEUE REGRESSIONSKOEFFIZIENTEN BERECHNEN
COMPUTE BETA=BETA+GRAD
** D. -2*LN LIKELIHOOD BERECHNEN
COMPUTE BEXP={ BETA, MAKE(K,1,0) }
COMPUTE PROB=(X*BEXP)
DO IF (MMAX(PROB) GT 170)
PRINT ITER /FOR "F2.0" /TITLE="++++ PROBLEM IN ITERATION:"
        /SPACE=5
PRINT /TITLE="EXP(X*BETA) NICHT DARSTELLBAR!"
        /SPACE=1
COMPUTE ERRMIN=1
BREAK
END IF
COMPUTE PROB=EXP(PROB)
COMPUTE TEMP=RSUM(PROB)*MAKE(1,I,1)
COMPUTE PROB=PROB &/TEMP
COMPUTE FIT=-2*MSUM((DES&*(LN(PROB))))
RELEASE TEMP
** E. ERGEBNISS IN ITERATIONSPROTOKOLL FESTHALTEN
COMPUTE KRIT=1-(FIT/FOLD)
COMPUTE LOG((ITER+1),2)=FIT            /* FUNKTIONSWERT FESTHALTEN */
COMPUTE LOG((ITER+1),3)=KRIT           /* KRITERIUM FESTHALTEN */
***  BEWERTUNG DER ERGEBNISSE
** A. FIT HAT SICH VERSCHLECHTERT
DO IF (KRIT < 0)
PRINT ITER /FOR "F2.0" /TITLE="++++ PROBLEM IN ITERATION:"
        /SPACE=5
PRINT /TITLE="     DEVIANZ HAT SICH ERHOEHT!"
        /SPACE=1
PRINT /TIT "+++++ PROGRAMM STOPPT MINIMIERUNG!"
COMPUTE ERRMIN=1
BREAK
END IF
** B. KONVERGENZ: FITZUWACHS < EPS1
DO IF (KRIT <= EPS1)
PRINT ITER /FOR "F2.0" /TITLE="***** KONVERGENZ NACH"
       /RLAB "ITERAT.:"
       /SPACE=5
COMPUTE CONV=1
END IF
** C. NAECHSTE ITERATION, WENN FITTZUWACHS > EPS1
DO IF (CONV=1)
BREAK
END IF
END LOOP       /* ENDE MINIMIERUNGSSCHLEIFE
```